Creating a Shared Reality with Robots

Faizan Muhammad, Amel Hassan, Andre Cleaver, and Jivko Sinapov
Email: {faizan.muhammad, amel.hassan, andre.cleaver, jivko.sinapov}@tufts.edu
Department of Computer Science
Tufts University, USA

***Abstract*—This paper outlines the system design, capabilities and potential applications of an Augmented Reality (AR) framework developed for Robot Operating System (ROS) powered robots. The goal of this framework is to enable high-level human-robot collaboration and interaction. It allows the users to visualize the robot's state in intuitive modalities overlaid onto the real world and interact with AR objects as a means of communication with the robot. Thereby creating a shared environment in which humans and robots can interact and collaborate.**

***Author Keywords*—Augmented Reality; Intelligent Robots; Human-Robot Interaction**

I. Introduction

As humans, there is a limit to what we can mentally process and physically execute. On the other hand, robots find it difficult to emulate traits like creative problem solving, empathy and intuition. Therefore, to efficiently tackle a particular scenario, an ideal human-robot collaboration system would capitalize on the specialized abilities of humans and robots to overcome their respective shortcomings. The key challenge here is that humans and robots use vastly different means of communication. While humans might employ speech, gestures and body language, robots usually rely on digital signalling. This presents difficulties for human teammates in understanding a robot's goals, intentions, knowledge and planning as it is working alongside them.

We believe Augmented Reality (AR) can be this required bridge from digital to analog and that it can be used to create a shared reality between humans and robots for communicating and problem-solving. To bring this shared reality to life, we have developed an AR framework that is capable of expressing the robot's sensory, planning, and cognitive information by projecting it visually in Augmented Reality. Furthermore, interaction with AR objects in this shared reality can be used as a basis for interaction with the robot at a high level.

We are exploring applications of this interface in many areas of human-robot collaboration and interaction, including education, navigation in shared spaces, and search and rescue. Our goal is to evaluate and improve this framework while identifying its limitations and further potential.

II. Prior Work and Motivation

Augmented Reality is an emerging technology that is increasingly being used in robotics. A significant amount of work has been done in exploring uses of AR in robotics such as education [1], target search [2], conveying robot motion intent [3] and teleoperation [4].

Our goal is to go a step further and build a generalized framework that can be easily adapted to any of these tasks. Moreover, we want this framework to leverage the robot's situational awareness, problem solving, and autonomy, thereby identifying the best modalities to express itself in AR and use that expression and its subsequent reaction from human users to solve problems.

III. Technical Details

We are using Robot Operating System (ROS) on the robot end, which is one of the most commonly used software architectures for robots. On the AR device end, we are using Unity as our base architecture given its versatility with devices ranging from iPads to Android phones to Microsoft Hololens. Figure 1 outlines the overall system architecture.

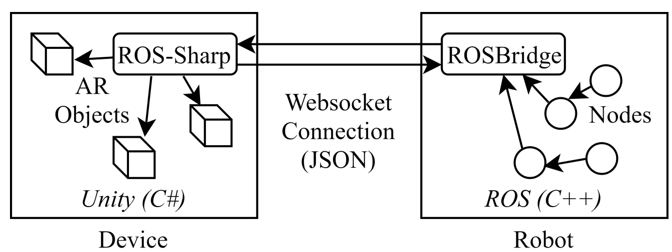


Fig. 1: System Architecture - AR objects are manifestations of ROS Topics

We are using a Websocket Connection between the device and the robot to allow ROSBridge to serialize topics within ROS and send them as JSON to Unity. Within Unity, we are using ROS-Sharp to convert these JSON messages back into C# data structures. We use a similar approach to communicate in the other direction. This gives us real-time, asynchronous and robust data communication within our system.

On the ROS end, we have written additional nodes to sample and compress regular ROS topics to save network bandwidth and prevent latency. Our nodes are also responsible for transforming all the coordinates in these topics to base_link frame of reference so that they can be projected in AR relative to the robot. We also have an ActionServer that is responsible for taking question/prompt requests from any ROS node in the robot and then forwarding them to Unity, getting a response and dispatching it back to the requester node.

This research has been conducted with funding support from the Tufts Center for Applied Brain and Cognitive Sciences (CABCS), and PTC Inc.

On Unity end, we have written scripts to visualize the robots data as objects and markers in AR. We also have scripts to interact with the robot in AR by publishing certain topics back into the robot. This gives the robot the ability to, for example, ask questions using visual prompts in AR. For tracking the robot relative to the user's AR device, we are using Vuforia and a custom Tracker Cube with high image feature density as observed in Figure 2. The high feature density of this cube is critical for getting precise localization of the robot relative to the device.

IV. Capabilities and Applications

The current capabilities of this AR framework can be divided into three categories. First, it allows the user to visualize the robot's sensory information to get an understanding of the robot's perspective of the world around it. Figure 2a is a screenshot depicting the visualization of the robot's laser scan. Visualizations like these can be very useful in education settings as they lower the expertise barrier required to be able to understand and debug aspects of robot's behavior that are based directly on sensory data [1].

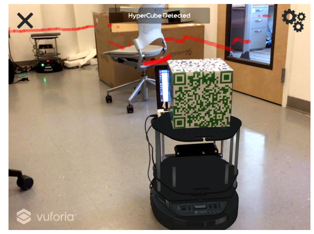
Second, the framework is able to visualize robot's planning and cognitive information. Figure 2b shows a visualization of the robot's intended path whereas Figure 2c shows the robot's local costmap. This is valuable in collaborative settings where it might be important to know what the robot is trying to do or about to do. It is also useful in explaining why a robot might be behaving a certain way. For example, by looking at the costmap, one can see why the robot picked a certain path as the safest. It is also useful in situations like search tasks, during which the robot can help find specific targets and share them with the user. For example, Figure 2d shows the visualization of the robot's People Detection.

Third, the interface allows the robot to prompt the user with textual notes, warnings, and general information as well as ask questions. While the previous two categories are non-invasive in terms of robot's behaviour (i.e. they only visualize it but do not, in any way, modify it), questions asked by the robot can be used to influence the robot's decision-making and expand its capabilities. For example, Figure 3 shows a robot requesting the user to open the door for it, which demonstrates its enhanced ability to interact with the world using this framework.

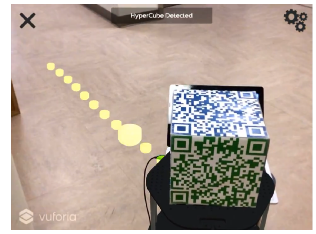
V. Future Work

We are currently working on two studies employing this framework. The first study investigates the use of different AR stimuli to avoid human-robot collisions in shared navigation spaces. Equipping social robots to convey their motion intent in social spaces is an ongoing problem and conventional signalling practices, such as LED indicators, are not extremely effective [5]. We hope that using stimuli in AR can further minimize collisions in shared spaces given that the stimuli can be designed to be highly sophisticated and adaptive.

The second study intends to investigate optimal modalities of data visualization in different stress environments during the task of search and rescue. It is intended to expand upon work done by Reardon et al [2], in which a robot provides navigational directions to a human user after locating a target. Our objective is to investigate how the effectiveness of different modalities for the navigational directions change at different induced stress levels in the human users.

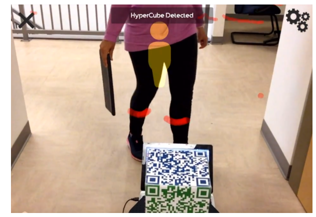


(a) Robot's Laser Scan seen as red particles



(b) Robot's Path Plan in AR seen as yellow dots

(c) Robot's Local Costmap seen as a grid



(d) Robot's People Detection seen as yellow icons

Fig. 2: Screenshots of the framework from an Apple iPad

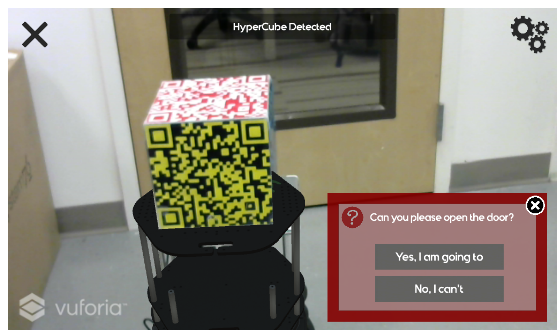


Fig. 3: Robot Asking the User a Question

References

[1] M. Cheli, J. Sinapov, E. Danahy, and C. Rogers, "Towards an augmented reality framework for K-12 robotics education," 1st International Workshop on Virtual, Augmented and Mixed Reality for Human-Robot Interaction (VAMHRI), 2018

[2] C. Reardon, K. Lee, and J. Fink, "Come See This! Augmented Reality to Enable Human-Robot Cooperative Search," 2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), 2018.

[3] M. Walker, H. Hedayati, J. Lee, and D. Szafir, Communicating Robot Motion Intent with Augmented Reality, Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction - HRI 18, 2018.

[4] H. Hedayati, M. Walker, and D. Szafir, Improving Collocated Robot Teleoperation with Augmented Reality, Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction - HRI 18, 2018

[5] R. Fernandez, N. John, S. Kirmani, J. Hart, J. Sinapov, and P. Stone, "Passive Demonstrations of Light-Based Robot Signals for Improved Human Interpretability," 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2018.